# Detecting non-trivial computation in complex dynamics

Joseph T. Lizier[1,2], Mikhail Prokopenko[1], and Albert Y. Zomaya[2]

[1] CSIRO Information and Communications Technology Centre, Locked Bag 17, North Ryde, NSW 1670, Australia
[2] School of Information Technologies, The University of Sydney, NSW 2006, Australia
jlizier@it.usyd.edu.au

**Abstract.** We quantify the local information dynamics at each spatiotemporal point in a complex system in terms of each element of computation: information storage, transfer and modification. Our formulation demonstrates that information modification (or non-trivial information processing) events can be locally identified where "the whole is greater than the sum of the parts". We apply these measures to cellular automata, providing the first quantitative evidence that collisions between particles therein are the dominant information modification events.

## 1 Introduction

Information-theoretic measures are increasingly being used to capture dynamics and to drive evolution in *artificial life*. Examples here include the use of a memory-like measure in [1], and information transfer-like measures in [2] and [3]. Such work appears disjointed however in that each example uses a different *single* measure of fitness. We observe that these single information-theoretic measures of fitness are often related to sub-functions or elements of computation. Interestingly, emergent complex behavior has often been described from the perspective of computation within the given system [4], and complex behavior is postulated to be associated with the capability to support universal computation [5, 6]. Such discussions focus on cellular automata (CAs) as model systems offering a range of dynamic behavior [4], including producing emergent structures such as gliders and glider collisions [7]. These discussions typically surround qualitative observation of the component operations of computation: *information storage*, *transfer* and *modification* (e.g. [5, 4]).

We suggest that a more intricate approach of *quantifying* the information dynamics of *each element* of computation will provide greater insight into and greater control over artificial life systems. We will describe how to quantify each element of computation on a local scale *within* a given system, showing how information storage and transfer interact to produce information modification; neither a single measure or system-wide approach is capable of this. We quantify a sum of the parts of computation, and locally identify information modification events where the sum is missing information; i.e. where the *whole is* *greater than the sum of the parts*. This phrase is often used to describe emergent structure in complex systems, e.g. patterns in Belousov-Zhabotinsky media [8], self-organization in microtubules [9] and collisions in CAs [4].

Our approach will provide insight into the local information dynamics of complex systems, from the perspective of the individual elements of computation. Here, we use it to demonstrate that the whole is quantitatively greater than the sum of the parts at collisions in CAs, and thereby prove the long-held conjecture that these are the dominant information modification agents therein.

## 2 Information-theoretical preliminaries

To quantify the elements of computation, we look to information theory (e.g. see [10]) which has proven to be a useful framework for the design and analysis of complex self-organized systems, e.g. [1–3]. The fundamental quantity is the *Shannon entropy*, which represents the uncertainty associated with any measurement $x$ of a random variable $X$ (logarithms are in base 2, giving units in bits): $H(X) = -\sum_x p(x) \log p(x)$. The *joint entropy* of two random variables $X$ and $Y$ is a generalization to quantify the uncertainty of their joint distribution: $H(X,Y) = -\sum_{x,y} p(x,y) \log p(x,y)$. The *conditional entropy* of $X$ given $Y$ is the average uncertainty that remains about $x$ when $y$ is known: $H(X|Y) = -\sum_{x,y} p(x,y) \log p(x|y)$. The *mutual information* between $X$ and $Y$ measures the average reduction in uncertainty about $x$ that results from learning the value of $y$, or vice versa: $I(X;Y) = H(X) - H(X|Y)$. The *conditional mutual information* between $X$ and $Y$ given $Z$ is the mutual information between $X$ and $Y$ when $Z$ is known: $I(X;Y|Z) = H(X|Z) - H(X|Y,Z)$.

The *entropy rate* is the limiting value of the conditional entropy of the next state $x$ of $X$ given knowledge of the previous $k-1$ states $x^{(k-1)}$ of $X$: $h_\mu = \lim_{k\to\infty} H\left[x|x^{(k-1)}\right] = \lim_{k\to\infty} h_\mu(k)$. Finally, the *excess entropy* quantifies the total amount of structure or memory in a system, and is computed in terms of the slowness of the approach of the entropy rate estimates to their limiting value (see [11]). For our purposes, it is best formulated as the mutual information between the semi-infinite past and semi-infinite future of the system:

$$E = \lim_{k\to\infty} I\left[x^{(k)}; x^{(k^+)}\right], \tag{1}$$

where $x^{(k^+)}$ refers to the next $k$ states. This interpretation is known as the *predictive information* [12], as it highlights that the excess entropy captures the information in a system's past which is relevant to predicting its future.

## 3 Introduction to Cellular Automata

Cellular automata (CA) are discrete dynamical systems consisting of an array of cells which each synchronously update their state as a function of the states of a fixed number of spatially neighboring cells using a uniform rule. While the behavior of each cell is simple, their (non-linear) interactions can lead to quite intricate global behavior. As such, CAs have become the classical example of complex behavior, and been used to model a wide variety of real world phenomena (see [4]). *Elementary CAs*, or *ECAs*, are a simple variety of 1D CAs using binary states, deterministic rules and one neighbor on either side (i.e. cell range $r = 1$). An example evolution of an ECA may be seen in Fig. 1a. For more complete definitions, including that of the Wolfram rule number convention for describing update rules (used here), see [13].

An important outcome of Wolfram's well-known attempt to classify the asymptotic behavior of CA rules into four classes [6, 13] was a focus on emergent structure: *particles*, *gliders* and *domains*. A domain is a set of background configurations in a CA, any of which will update to another configuration in the set in the absence of any disturbance. A domain may be *regular*, where the configurations repeat periodically, or is otherwise known as *irregular*. Domains are formally defined within the framework of computational mechanics as spatial process languages in the CA [14]. Particles are considered to be dynamic elements of coherent spatiotemporal structure; gliders are regular particles, blinkers are stationary gliders. Formally, particles are defined by computational mechanics as a boundary between two domains [14]; they can be termed *domain walls*, though this is typically used with reference to irregular domains. Several techniques exist to *filter* particles from background domains (e.g. [15, 16]).

## 4 Computation in Cellular Automata

Computation in CAs has been a popular topic for study, with a major focus in observing or constructing (Turing) universal computation in certain CAs (see [4–6]). This capability has been proven for several CA rules (e.g. the Game of Life [7]), through the design or identification of entities which provide the three primitive functions of universal computation: information storage, transmission and modification. Typically such analyses focus on blinkers as the basis of information storage, particles as the basis of information transfer, and collisions between these structures as information modification (see [5, 4]).

However, the focus on universal computational ability has been criticized as drawing away from the ability to identify "generic computational properties" in these and other CAs [14]. Related criticisms target attempts to classify CA rules in terms of generic behavior or "bulk statistical properties", suggesting that the wide range of dynamics taking place in different areas of the CA make this problematic [14, 4]. With respect to computation, it would be too simplistic to say a CA was either *computing* or *not computing*. Alternatively, these studies suggest that analyzing the rich space-time dynamics *within* the CA is a more appropriate focus, since the CA may be undertaking different parts of a complex computation at different times or spatial points. As such, these and other studies have *analyzed the local dynamics of intrinsic or other specific computation*, while continuing the focus on particles facilitating the transfer of information and collisions facilitating information modification or processing. Noteworthy examples include: applying filters from the domain of computational mechanics (using regular language parsers in [14] and local statistical complexity in [15]); analysis using such filters on CA rules selected via evolutionary computation to perform tasks including classification [17]; and deep investigations of particle properties and their interactions [17, 18].

Despite this surrounding interest, no complete local quantification of the individual elements of computation exists. In the subsequent sections, we outline how the individual elements of computation can be locally quantified within the spatiotemporal structure of a CA. In particular, we describe how information storage and information transfer interact to give rise to information modification events, being where the whole is greater than the sum of the parts.

## 5 Information Storage

Although discussion of information storage, or memory, in CAs has often focused on periodic structures, it does not necessarily entail periodicity. Instead, the excess entropy (1) encompasses all types of structure and memory by capturing correlations across all time lengths. In examining *local* information dynamics, we are particularly interested in how much of the stored information is actually *in use* at every local point in time and space.

The excess entropy can be massaged into a spatiotemporally local measure (i.e. the amount of information stored by a *particular* cell at a *particular* point in time) by noting that it is actually the *expectation value* of a local excess entropy at every time step [19].[3] The local excess entropy is then the mutual information between the semi-infinite past and future *for the given cell at the given time step*. It quantifies the total stored information that will be used *at some point* in the future of the state process of that cell; possibly but not necessarily at the next time step $n+1$. To reveal the amount of memory actually *in use* at the next time step, we derive *local active information storage* $a(i, n+1)$ as the local mutual information between the semi-infinite past $x_{i,n}^{(k)}$ (as $k \rightarrow \infty$) and the *next state* $x_{i,n+1}$ of a given cell $i$ at the given time step $n+1$:

$$a(i, n+1) = \lim_{k \rightarrow \infty} \log \frac{p(x_{i,n}^{(k)}, x_{i,n+1})}{p(x_{i,n}^{(k)}) p(x_{i,n+1})}. \tag{2}$$

It is not feasible to compute $a(i, n)$ in the limit $k \rightarrow \infty$; instead we compute $a(i, n, k)$ with finite $k$. Importantly, $a(i, n, k)$ may be positive or negative, meaning the past history of the cell can either positively inform us or actually *misinform* us about it's next state. An observer is misinformed where, given the past history, the observed outcome was relatively unlikely.

[3] As per Shalizi's explanation in [19], which was for the *light-cone formulation* of excess entropy. A detailed description on why such average measures are the expectation value of local measures, and why the local measure is simply the log term within the expectation value, lies in the presentation of local transfer entropy in [16].

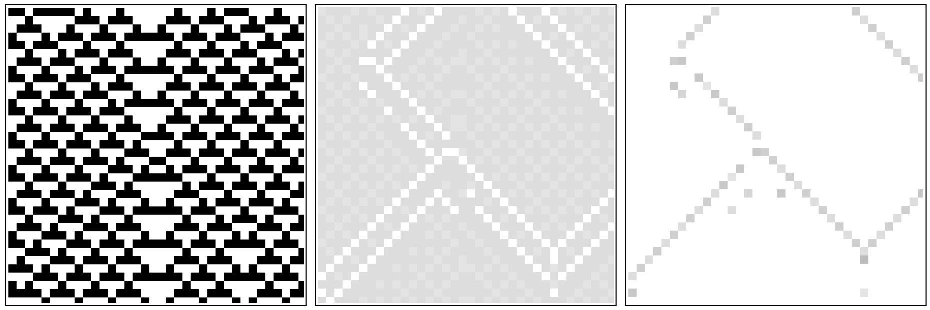

**Fig. 1. Information Storage**. ECA Rule 54: a. (*left*) Raw CA (time is vertical). b.,c. Local active information storage $a(i, n, k = 16)$: b. (*center*) positive values only, grayscale (30 levels), max. 1.11 bits (black); c. (*right*) negative values only, grayscale (30 levels), min. -12.2 bits (black). All figures generated using modifications to [20].

In order to evaluate $a(i, n, k)$ (and our later metrics) within sample CA runs, we estimate the probability distribution functions from CA runs of 10 000 cells with periodic boundary conditions, initialized from random states, with 600 time steps retained (after the first 30 time steps were eliminated to allow the CA to settle). Since the CA is a homogeneous system, the probability functions were estimated from observations at every spatiotemporal point so as to obtain more accurate estimates. All results were confirmed by at least 10 runs from different initial states. We use ECA rule 54 as a running case study: it is complex enough to contain traveling and stationary gliders, though these are simple enough that the results are easily understood.

The raw states of a sample run for rule 54 are displayed in Fig. 1a. The local active information storage $a(i, n, k = 16)$ profile contains both positive and negative values, plotted separately in Fig. 1b and c. The positive values are concentrated in the domain areas, as well as in the stationary gliders (known as $\alpha$ and $\beta$ [18]), due to strong temporal periodicity in both areas. In fact, the $a(i, n, k)$ values form spatially and temporally periodic patterns in the domain, due to the periodicities of the raw states. The negative values are concentrated in the traveling glider areas (known as $\gamma^+$ and $\gamma^-$ [18]). When a traveling glider is encountered at a given cell, the past history of that cell (being part of the background regular domain) is misinformative about the next state of the cell, since the domain sequence was more likely to continue than be interrupted. It is useful that $a(i, n, k)$ can reveal the gliders in this way (similar to existing filtering methods for CAs [14, 15]), however it is does not tell us the source of information for the glider: to quantify this, we turn to information transfer.

## 6 Information Transfer

Schreiber presented *transfer entropy* as a directional, dynamic measure for information transfer [21] so as to address deficiencies in the de facto use of mutual information (as a static, symmetric measure). Transfer entropy is the deviation from independence (in bits) of the state transition of an information destination $X$ from the (previous) state of a source $Y$. It is a *conditional* mutual information [22], being the average information in the source $y_n$ about the next state of the destination $x_{n+1}$ that was not already contained in the destination's past $x_n^{(k)}$ (i.e. $I(Y_n; X_{n+1}|X_n^{(k)})$), allowing us to quantify directional flows of information.

In [16], we demonstrated that the transfer entropy is an *expectation value* of a *local* transfer entropy at each observation. We also generalized comments on the entropy rate in [21] to suggest that the asymptote $k \to \infty$ is most correct for agents displaying non-Markovian dynamics. For systems such as CAs with homogeneous spatially-ordered agents, the local apparent transfer entropy to cell $X_i$ from $X_{i-j}$ at time $n+1$ is:

$$t(i, j, n+1) = \lim_{k \to \infty} \log \frac{p(x_{i,n+1}|x_{i,n}^{(k)}, x_{i-j,n})}{p(x_{i,n+1}|x_{i,n}^{(k)})}, \qquad (3)$$

for transfer from the previous time step only. Again, we define $t(i, j, n, k)$ for finite $k$. Transfer $t(i, j, n, k)$ is defined for every spatiotemporal destination $(i, n)$, for every information channel or direction $j$ where sensible values for CAs are within the cell range, $|j| \leq r$ (e.g. $j = 1$ means transfer across one cell to the right). Local apparent transfer entropy $t(i, j, n, k)$ may be either positive or negative, with negative values occurring where (given the destination's history) the source is actually *misleading* about the next state of the destination.

The destination's own historical values can indirectly influence it via the source or other neighbors and be mistaken as an independent flow from the source [16]. In the context of computation, this influence is recognizable as the *active information storage. The active information storage* $a(i, n+1)$ *is eliminated from the transfer entropy measurement by conditioning on the destination's history* $x_{i,n}^{(k)}$. Yet any self-influence transmitted prior to these $k$ values will not be eliminated, which is why we suggest taking the limit $k \to \infty$ to be most correct.

We applied the local transfer entropy metric to several important ECA rules in [16]. Fig. 2 displays application of the local apparent transfer entropies to rule 54, demonstrating that the metric successfully highlights traveling gliders with large positive transfer against background domains (it also highlights domain walls where they exist). Importantly, the metric finds negative transfer for gliders moving orthogonal to the direction of measurement, because the source (as part of the domain) is misinformative about the next state of the destination. Also, there is a small non-zero information transfer in background domains, effectively indicating the absence of gliders; this is particularly strong in the wake of real gliders, where secondary gliders often follow. *The measure provided the first quantitative evidence for the long-held conjecture that particles are the dominant information transfer agents in CAs.* This highlighting was similar to other methods of filtering in CAs (e.g. [15, 14]), but subtly allowed comparison between and within gliders of the amount and (channel or) direction of information transfered at each point, and revealed the leading edges of gliders as the

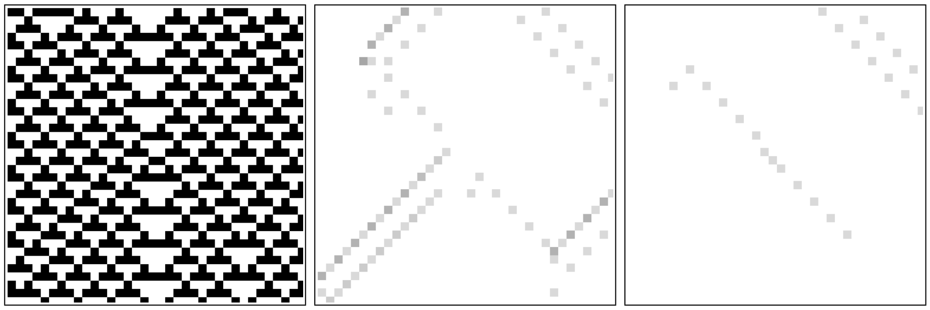

**Fig. 2. Information Transfer**. ECA Rule 54: a. (*left*) Raw CA. b.,c. Local apparent transfer entropy $t(i, j = -1, n, k = 16)$ ($j = -1$ means transfer one cell to the left): b. (*center*) positive values only, grayscale (16 levels), max. 7.92 bits (black); c. (*right*) negative values only, grayscale (16 levels), min. -4.21 bits (black).

major information transfer zones. At least a minimum $k$ was required to achieve reasonable estimates of the metric (e.g. of the order of the period of a regular periodic domain); without this, particles were not highlighted. Finally, note that this metric cannot quantitatively distinguish gliders from their collisions: for this, we look to an information modification metric.

## 7 Information Modification

Information modification has been described as interactions between stored and transmitted information that result in a modification of one or the other [5], and generally interpreted to mean interactions or collisions of particles. As an information processing event, the important role of collisions in determining the dynamics of the system is widely acknowledged [18]. For a regular particle or glider, a local *modification* is simply an alteration to the predictable periodic pattern of the glider's dynamics, where an observer would be surprised or misinformed about the next state of the glider without having taken account of the entity about to perturb it. Recall that local apparent transfer entropy $t(i, j, n)$ and local active information storage $a(i, n)$ were negative where the respective information sources were *misinformative* about the next state of the information destination. This occurred for $a(i, n)$ at unperturbed gliders, and for $t(i, j, n)$ at gliders traveling in the orthogonal direction to the measurement. However, we expect that the $t(i, j, n)$ in the direction of motion of the glider will be *more informative* than the misinformation conveyed from the other sources.

Where a glider is perturbed by an interaction with another glider, we cannot expect $t(i, j, n)$ in the macroscopic direction of the first glider to remain informative about the evolution of that glider at the collision point. Nor can we expect this from the incoming $t(i, j, n)$ for the incident glider. As such, we hypothesize that *at the spatiotemporal location of a local information modification event or collision, the total information from the information storage and information* *transfer, where each source is observed separately, will misinform an observer.* We label this total as the *local separable information*, $s(i, n)$:

$$s(i, n) = a(i, n) + \sum_{j=-r, j \neq 0}^{+r} t(i, j, n), \tag{4}$$

with $s(i, n, k)$ representing the approximation for finite $k$. Where $s(i, n)$ is *positive* or *highly separable*, separate or independent observations of the sources are informative overall about the next state of the information destination. This indicates that information storage and transfer are not interacting, and only trivial information modifications are taking place. Conversely, we expect $s(i, n)$ to be *negative* at spatiotemporal points where an information modification event or collision takes place, with more significant modifications taking larger negative values. Separate examination of sources fails here because the information storage and transfer are interacting, i.e. *non-trivial information modification* takes place. This formulation of non-trivial information modification quantifies the description of emergence in complex systems as where *"the whole is greater than the sum of it's parts"*. While we quantify the sum of the parts in $s(i, n)$, there is no quantity representing the "whole" as such, simply an indication that the whole is greater where all information sources must be examined *together* in order to receive positive information on the next state of the given entity.

Fig. 3 displays application of $s(i, n, k)$ to ECA rule 54. Positive values of $s(i, n, k)$ (not plotted) are concentrated in the domain regions and at the stationary gliders ($\alpha$ and $\beta$): as expected, these regions are undertaking trivial computations only. The dominant negative values of $s(i, n, k)$ are concentrated around the areas of collisions between the gliders, including those between traveling gliders only (marked by "A") and between the traveling gliders and blinkers (marked by "B" and "C"). Considering the collision "A" ($\gamma^{+} + \gamma^{-} \rightarrow \beta$ [18]), the marked information modification event is one time step below where one may naively define it. Our metric correctly marks the information modification event however, being where prediction requires the sources to be considered together. For the other collisions "B" and "C" also, the spatiotemporal location of the primary information modification(s) appears to be delayed from a naively defined collision point; this indicates a time-lag associated with processing the information. Smaller negative values are also associated with the gliders (too small to appear in Fig. 3b), which was unexpected. These weak information modifications appear to indicate the absence of a collision (i.e. the absence of an incident glider) and in some sense are a computation that the glider will continue. These computations are more significant in the wake of real collisions (indeed the secondary collision points for types "B" and "C" are higher in magnitude than the earlier collision points), since they have a larger influence on the surrounding dynamics at those points. This finding is analogous to that of small values of transfer entropy in domains indicating the absence of gliders, which were also more significant in the wake of real gliders [16]. This is the first known metric which brings together information storage and transfer to identify information modification, and it has provided the first quantitative evidence that collisions

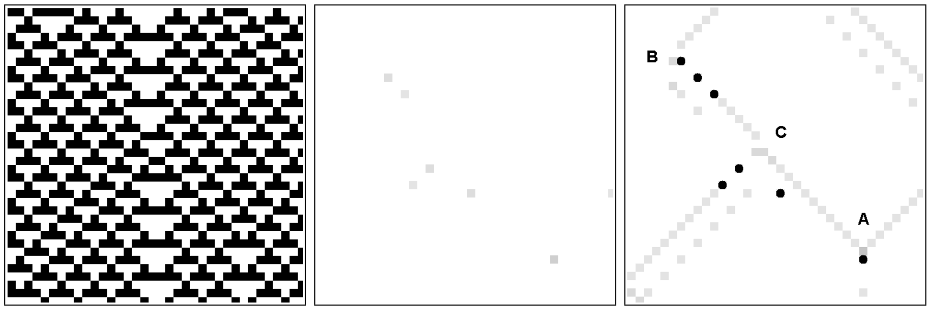


**Fig. 3. Information Modification**. ECA Rule 54: a. (*left*) Raw CA. b. (*center*) Local separable information $s(i, n, k = 16)$, negative values only plotted, grayscale (30 levels), min -5.23 bits (black). c. (*right*) Locations of negative values of $s(i, n, k = 16)$ (larger than weak values along the gliders) marked with black circles against $t(i, j, n, k = 16)$ summed over $j = [-1, 1]$; "A", "B" and "C" mark collision types discussed in the text.

in CAs are the dominant information modification events therein. It is also the first suite of filters able to distinguish between particles and particle collisions.

## 8 Conclusion

We have discussed appropriate local metrics for information storage and transfer, and demonstrated how these quantities interact to produce information modification events where the "whole is greater than the sum of the parts". These metrics form a powerful framework for quantifying the information dynamics of complex systems. Here, the framework has been applied to CAs, providing the first evidence for the long-held conjecture that collisions therein are the dominant information modification events.

We aim to provide a deeper investigation of this topic in a forthcoming analysis, reporting results we have obtained from application of the presented methods to other CA rules (complex, chaotic and those with domain walls) which corroborate and extend those reported here. We believe these three elements of computation identify a set of *axes of complexity* to characterize complex systems. Within this multi-dimensional space, a system such as a CA may appear complex in one dimension (e.g. information storage) without necessarily appearing so in the other dimensions. Finally, we shall explore the generality afforded by the information-theoretic basis of this framework. Application to higher dimensional CAs or other discrete dynamical systems is straightforward, and we expect application to artificial life systems to provide greater insight into and control over such systems.

## References

1. Prokopenko, M., Gerasimov, V., Tanev, I.: Evolving spatiotemporal coordination in a modular robotic system. In Nolfi, S., Baldassarre, G., Calabretta, R., Hallam,

J., Marocco, D., Meyer, J.A., Parisi, D., eds.: Ninth International Conference on the Simulation of Adaptive Behavior (SAB'06). Volume 4095 of Lecture Notes in Artificial Intelligence., Rome, Springer Verlag (2006) 548–559
2. Klyubin, A.S., Polani, D., Nehaniv, C.L.: All else being equal be empowered. In Capcarrere, M.S., Freitas, A.A., Bentley, P.J., Johnson, C.G., Timmis, J., eds.: 8th European Conference on Artificial Life (ECAL 2005). Volume 3630 of Lecture Notes in Computer Science., Kent, UK, Springer Berlin/Heidelberg (2005) 744–753
3. Lungarella, M., Sporns, O.: Mapping information flow in sensorimotor networks. PLoS Computational Biology **2**(10) (2006) e144
4. Mitchell, M.: Computation in cellular automata: A selected review. In Gramss, T., Bornholdt, S., Gross, M., Mitchell, M., Pellizzari, T., eds.: Non-Standard Computation. VCH Verlagsgesellschaft, Weinheim (1998) 95–140
5. Langton, C.G.: Computation at the edge of chaos: phase transitions and emergent computation. Physica (Amsterdam) **42D**(1-3) (1990) 12–37
6. Wolfram, S.: Universality and complexity in cellular automata. Physica (Amsterdam) **10D**(1-2) (1984) 1–35
7. Conway, J.H.: What is life? In Berlekamp, E., Conway, J.H., Guy, R., eds.: Winning ways for your mathematical plays. Volume 2, ch. 25. Academic Press, New York (1982)
8. Sendiña-Nadal, I., Mihaliuk, E., Wang, J., Pérez-Muñuzuri, V., Showalter, K.: Wave propagation in subexcitable media with periodically modulated excitability. Phys. Rev. Lett. **86**(8) (2001) 1646
9. Brown, J.A., Tuszynski, J.A.: A review of the ferroelectric model of microtubules. Ferroelectrics **220** (1999) 141–156
10. MacKay, D.J.: Information Theory, Inference, and Learning Algorithms. Cambridge University Press, Cambridge (2003)
11. Crutchfield, J.P., Feldman, D.P.: Regularities unseen, randomness observed: Levels of entropy convergence. Chaos **13**(1) (2003) 25–54
12. Bialek, W., Nemenman, I., Tishby, N.: Complexity through nonextensivity. Physica (Amsterdam) **302A**(1-4) (2001) 89–99
13. Wolfram, S.: A New Kind of Science. Wolfram Media, Champaign, IL, USA (2002)
14. Hanson, J.E., Crutchfield, J.P.: The attractor-basin portait of a cellular automaton. J. Stat. Phys. **66** (1992) 1415–1462
15. Shalizi, C.R., Haslinger, R., Rouquier, J.B., Klinkner, K.L., Moore, C.: Automatic filters for the detection of coherent structure in spatiotemporal systems. Phys. Rev. E **73**(3) (2006) 036104
16. Lizier, J.T., Prokopenko, M., Zomaya, A.Y.: Local information transfer as a spatiotemporal filter for complex systems. Unpublished (2007)
17. Mitchell, M., Crutchfield, J.P., Hraber, P.T.: Evolving cellular automata to perform computations: Mechanisms and impediments. Physica (Amsterdam) **75D** (1994) 361–391
18. Hordijk, W., Shalizi, C.R., Crutchfield, J.P.: Upper bound on the products of particle interactions in cellular automata. Physica (Amsterdam) **154D**(3-4) (2001) 240–258
19. Shalizi, C.R.: Causal Architecture, Complexity and Self-Organization in Time Series and Cellular Automata. PhD thesis, University of Wisconsin-Madison (2001)
20. Wójtowicz, M.: Java cellebration v.1.50 (2002) Online Software.
21. Schreiber, T.: Measuring information transfer. Phys. Rev. Lett. **85**(2) (2000) 461–464
22. Ay, N., Polani, D.: Information flows in causal networks. Adv. Comp. Sys. (to be published) (2006)